Europ. J. Combinatorics (1995) **16**, 427–438

Forbidden Configurations: Induction and Linear Algebra

R. P. ANSTEE

Let a matrix be called simple if it is a (0, 1)-matrix with no repeated columns. We consider results of the form: if A is an $m \times n$ simple matrix with no submatrix which is a column permutation of F for all F in some specified set of matrices $\mathscr{F}$, then $n \leqslant f(m)$. We obtain some results using induction and some results using linear algebra.

1. INTRODUCTION

This paper considers some extremal problems of matrices with forbidden substructures. The following basic result gives the flavour. Define a matrix to be *simple* if it is a (0, 1)-matrix with no repeated columns (matrices in this paper are all (0, 1)-matrices). Let K_k be a $k \times 2^k$ simple matrix of all possible columns on k rows.

THEOREM 1.1 (Vapnik and Chervonenkis [10], Sauer [8] and Perles and Shelah [9]). *Let A be an $m \times n$ simple matrix with no submatrix which is a row and column permutation of K_k. Then*

$$n \leqslant \binom{m}{k-1} + \binom{m}{k-2} + \cdots + \binom{m}{0}. \tag{1.1}$$

This bound can be seen to be exact in that we can form a matrix of all columns of $k-1$ or fewer 1's to achieve equality in (1.1). More generally, for any $k \times 1$ (0, 1)-column α, the matrix of all columns with no submatrix α also achieves the bound (1.1) (Theorem 2.4 [4]).

Combinatorial objects are often represented by (0, 1)-matrices and when we forbid a matrix F we would usually forbid any row and column permutation of F. In this paper, we drop the row permutations since it is natural in both our proof techniques. If we wish to forbid any row and column permutation of F as a submatrix then we could just forbid any column permutation of any $F' \in \mathscr{F}$ as a submatrix, where $\mathscr{F}$ consists of all row permutations of F. Note that a row permutation of K_k is a column permutation of K_k, so that in Theorem 1.1 this distinction does not matter.

Our typical problem will start with a set $\mathscr{F}$ of matrices. Let A be a simple $m \times n$ matrix with no column permutation of F as a submatrix for each $F \in \mathscr{F}$. We ask for a bound on n in terms of m; i.e. a function $f: \mathbf{N} \to \mathbf{N}$ with $n \leqslant f(m)$. We also seek structural information about A.

Section 2 gives two induction arguments for certain structured $\mathscr{F}$. Section 3 considers the role of linear algebra in obtaining the bounds f as well as some structural information you can recover. If A is an $m \times n$ simple matrix and γ is an $m \times 1$ (0, 1)-column, then let $A(\gamma)$ denote the $1 \times n$(0, 1)-row with a 1 in precisely those columns β of A satisfying $\beta \geqslant \gamma$. Thus $A(\vec{0}_m)$ (where $\vec{0}_m$ is the column of m 0's is the $1 \times n$ vector of 1's and hence n is encoded by $A(\vec{0}_m)$. Now the vector space

$$\langle\{A(\gamma) \mid \gamma\, m \times 1\ (0,1)\text{-}columm\}\rangle \tag{1.2}$$

has rank over $\mathbf{Q}$ at most $f(m)$, assuming $n \leqslant f(m)$.

0195-6698/95/050427 + 12 $12.00/0

REMARK 1.2. $\{A(\gamma) \mid \gamma \text{ is a column of } A\}$ forms a basis for (1.2.)

PROOF. If you order the columns of A respecting the column partial order '$\leq$', the set of vectors $A(\gamma)$ can be seen to form a triangular system. □

What we do is find a set $\mathcal{F}_{[m]}$ of columns, determined by the forbidden submatrices $\mathcal{F}$ (not by A), with $|\mathcal{F}_{[m]}| = f(m)$, so that

$$\{A(\gamma) \mid \gamma \in \mathcal{F}_{[m]}\} \tag{1.3}$$

forms a spanning set for (1.2). Then this would prove the bound $n \leq f(m)$. We generalize the linear algebra proof of Theorem 1.1 [2, 5]. We are using the notation $[m] = \{1, 2, \ldots, m\}$ in $\mathcal{F}_{[m]}$ to indicate the rows on which the columns of $\mathcal{F}_{[m]}$ are given.

For a $m \times n$ matrix A let $t \cdot A$ denote the $m \times tn$ matrix of t copies of A. For an $m \times n$ matrix A, a $t \times 1$ column α, and a $(m + t) \times 1$ column β of m 1's and t 0's, define *merge*(A, α, β) as the $(m + t) \times n$ matrix obtained by placing the rows of A in order in the rows indexed by 1's of β and the rows of $n \cdot \alpha$ in the rows indexed by 0's of β. In Section 2 we obtain a bound for merge$(K_k, \vec{0}_l, \beta)$ (with β having k 1's and l 0's) where the bound is achieved by the columns avoiding the submatrix $\beta = \text{merge}(\vec{1}_k, \vec{0}_l, \beta)$. The linear algebra shows that every other matrix avoiding merge$(K_k, \vec{0}_l, \beta)$ and achieving the bound 'covers' the columns just given.

Section 4 obtains the bound from forbidding merge(K_k, α, β) for all $l \times 1$ (0,1)-columns α with β as above. A sporadic new result is also obtained.

Let E_k (respectively O_k) denote the $k \times 2^{k-1}$ simple matrix of all columns of even (resp. odd) column sum. For $S \subseteq [m]$ and A on m rows, let $A|_S$ denote the submatrix of A consisting of those rows of A indexed by S. Finally, since we are concerned with column permutations, then for two matrices A,B write $A =_c B$ if A is a column permutation of B.

2. INDUCTION

The following two propositions provide general induction tools for forbidden submatrices.

PROPOSITION 2.1. *Let $\mathcal{F}$ be a set of matrices and assume that there is a function $f: \mathbf{N} \to \mathbf{N}$ such that, for any i, if A is an $i \times n$ simple matrix with no submatrix which is a column permutation of an $F \in \mathcal{F}$, then $n \leq f(i)$. Then if B is an $m \times n$ simple matrix with no submatrix which is a column permutation of*

$$\bar{F} = \begin{bmatrix} 11 \cdots 1 & 00 \cdots 0 \\ F & F \end{bmatrix} \tag{2.1}$$

for any $F \in \mathcal{F}$, then

$$n \leq 2 + \sum_{i=1}^{m-1} f(i). \tag{2.2}$$

PROOF. Use induction on m. The base case $m = 1$ yields an upper bound of 2. Let B be an $m \times n$ simple matrix as described and partition B:

$$B =_c \begin{bmatrix} 11 \cdots 1 & 00 \cdots 1 \\ B_1 B_2 & B_2 B_3 \end{bmatrix}, \tag{2.3}$$

where B_2 consists of one copy of each repeated column in the matrix obtained from B

by deleting the first row. Thus $[B_1B_2B_3]$ is simple with no column permutation of $\bar{F}$ as a submatrix for every $F \in \mathcal{F}$ and by induction has at most

$$2+\sum_{i=1}^{m-2} f(i) \tag{2.4}$$

columns. Also B_2 is simple and has no column permutation of $\bar{F}$ as a submatrix for $F \in \mathcal{F}$; otherwise, B has $\bar{F}$ and so B_2 has at most $f(m-1)$ columns. Adding $f(m-1)$ to (2.4) yields (2.2). □

PROOF OF THEOREM 1.1 Use induction on k. Take $\mathcal{F}=\{K_{k-1}\}$ and note that $\bar{K}_{k-1}=_c K_k$. □

PROPOSITION 2.2. *Let $\mathcal{F}$ be a set of matrices and assume that there is a function $f: \mathbf{N}\to\mathbf{N}$ such that for any i, if A is an $i \times n$ simple matrix with no submatrix which is a column permutation of an $F \in \mathcal{F}$, then $n \le f(i)$. Then if B is an $m \times n$ simple matrix with no submatrix which is a column permutation of*

$$F'=\begin{bmatrix} 00\cdots 0 \\ F \end{bmatrix} \quad or \quad F''=\begin{bmatrix} 11\cdots 1 \\ F \end{bmatrix} \tag{2.5}$$

for any $F \in \mathcal{F}$, then $n \le 2\cdot f(m-1)$.

PROOF. Decompose B as in (2.3). Then $[B_1B_2]$ is simple and has no $F \in \mathcal{F}$ and so has at most $f(m-1)$ columns. Similarly, $[B_2B_3]$ has at most $f(m-1)$ columns, yielding the bound on n. □

A number of results can be generated by these tools. The following will be re-examined in Section 4. The definition of merge(K_k, α, β) is in Section 1.

COROLLARY 2.3. *Let k, l and a $(k+l)\times 1$ (0, 1)-vector β of exactly k 1's be given. Let A be an $m \times n$ simple matrix with no column permutations of* merge(K_k, α, β) *as a submatrix for any $l \times 1$ (0, 1)-column α. Then*

$$n \le 2^l\left(\binom{m-l}{k-1}+\binom{m-l}{k-2}+\cdots+\binom{m-l}{0}\right). \tag{2.6}$$

and this bound is best possible.

PROOF. Apply Propositions 2.1 and 2.2 repeatedly. If $\beta^{\mathrm{T}}=(b_{k+1}, b_k, \ldots, b_2, b_1)$ then at step i we use Proposition 2.1 if $b_i=1$ and Proposition 2.2 if $b_i=0$. We would start the process with $\mathcal{F}$ consisting of a single matrix of 1 column and 0 rows and $f(i)=0$ for all i. The bound is achieved by taking all columns with no $(k+l)\times 1$ submatrix merge $(\vec{1}_k, \alpha, \beta)$ as a submatrix for all $l \times 1$ α. This is analyzed in detail at the end of the proof of Theorem 4.3. □

An interesting construction shows the bound (2.6) to be best possible for $l=1$ even if we forbid row *and* column permutations of merge(K_k, α, β) as a submatrix.

REMARK. 2.4.There is an $m \times 2(\binom{m-1}{k-1}+\binom{m-1}{k-2}+\cdots+\binom{m-1}{0})$ simple matrix A with no row and column permutation of merge$(K_k, 0, \beta)$ or merge$(K_k, 1, \beta)$, where β is a $(k\times 1)\times 1$ column of k 1's and one 0.

PROOF. Take all $m \times 1$ columns with no $(k+1) \times 1$ submatrices $(1,0,1,0,\ldots)^T$ or $(0,1,0,1,\ldots)^T$, which is achieved by taking all $(m-1) \times 1$ columns with no $k \times 1$ submatrix $(0,1,0,1,\ldots)^T$ and inserting a 1 as row 1 and also the (0, 1)-complements of these columns. This yields the desired matrix, since any row and column permutation of merge$(K_k, 0, \beta)$ or merge$(K_k, 1, \beta)$ has either $(1,0,1,0,\ldots)^T$ or $(0,1,0,1,\ldots)^T$ as $(k+1) \times 1$ submatrices. □

For $\mathcal{F} = \{\vec{1}_k, \vec{0}_k\}$ we obtain the bound $f(i) = 0$ for $i \geq 2k-1$. Applying Proposition 2.1 repeatedly we obtain the following:

COROLLARY 2.5. *Let β be the $(k+l) \times 1$ (0, 1)-columns with k 1's in the top k rows. Let A be an $m \times n$ simple matrix with no column permutation of merge$(K_k, \vec{0}_l, \beta)$. Then n is $O(m^{k-1})$ and this is asymptotically best possible.*

COROLLARY 2.6 (Füredi [6]). *Let A be an $m \times n$ simple matrix with no column permutation of $t \cdot K_k$. Then n is $O(m^k)$.*

PROOF. Simply note that the bound for $\underline{t \cdot K_1}$ is $O(n)$ (Theorem 4.3 [4]) and then use induction on k and Proposition 2.1 with $\overline{t \cdot K_{k-1}} = t \cdot K_k$. □

PROPOSITION 2.7. *Let $\mathcal{F}$ consist of all $(k+l) \times 1$ (0, 1)-columns with k* 1's and l 0's. *Let A be an $m \times n$ simple matrix with no submatrix $F \in \mathcal{F}$. Then*

$$n \leq \binom{m}{k-1} + \binom{m}{k-2} + \cdots + \binom{m}{0} + \binom{m}{l-1} + \binom{m}{l-1} + \cdots + \binom{m}{0}.$$

PROOF. Columns avoiding $F \in \mathcal{F}$ must have fewer than k 1's or fewer than l 0's. □

COROLLARY 2.8. *Let A be an $m \times n$ simple matrix with no column permutation of* merge(K_t, α, β) *as a submatrix for $\alpha \in \mathcal{F}$ as given in proposition 2.7 and β being the $(t+k+1) \times 1$ (0, 1)-column with t 1's in the first t rows. Assume $k \geq l$. Then*

$$n \leq (1 + \delta_{k,l})\binom{m}{t+k-1} + O(m^{t+k-2}),$$

where $\delta_{k,l}$ is 1 *if $k = 1$ and* 0 *otherwise.*

PROOF. Apply Proposition 2.1 t times to Proposition 2.7, focusing on the leading terms. □

3. LINEAR ALGEBRA

In what follows we give a series of results that give linear algebra proofs of the bounds obtained by forbidding any column permutation of $F \in \mathcal{F}$ as a submatrix as well as some structural results. We will generalize the linear algebra proof of Theorem 1.1 of Frankl and Pach [5] and independently [2]. The notation will be used again in Section 4 and so is more general than required here.

Let α be a given $(k \times l) \times 1$ (0, 1)-column with k 1's and l 0's. Define

$$\mathcal{F}_{[m]} = \{m \times 1 (0, 1)\text{-column } \gamma \mid \gamma \text{ has no submatrix } \alpha]. \tag{3.1}$$

We note as in [4] that

$$|\mathcal{F}_{[m]}| = \binom{m}{k+l-1} + \binom{m}{k+l-2} + \cdots + \binom{m}{0}, \tag{3.2}$$

since if $\alpha^{\mathrm{T}} = (a_1, a_2, \ldots, a_{k+l})$ then, for any $\gamma \in \mathcal{F}_{[m]}$, γ^{T} is an initial segment of $\bar{a}_1^* a_1 \bar{a}_2^* \cdots \bar{a}_{k+l-1}^* a_{k+l-1} \bar{a}_{k+l}^*$, where $\bar{0} = 1$, $\bar{1} = 0$ and a^* denotes an arbitrary sequence of a's, possibly empty. Let

$$F_1' = \mathrm{merge}(E_k, \vec{0}_l, \alpha), \qquad F_1'' = \mathrm{merge}(O_k, \vec{0}_l, \alpha). \tag{3.3}$$

We set $t = 1$, $s = k + l$ and let

$$\mathcal{F} = \{[F_i' F_i''] \mid i = 1, 2, \ldots, t\}, \tag{3.4}$$

so in our case, $\mathcal{F} = \{\mathrm{merge}(K_k, \vec{0}_l, \alpha)\}$. Let $\#_1(x)$ be the number of 1's in x.

THEOREM 3.1. *Let A, B be two $(0, 1)$-matrices on m rows with*

$$\#_1(A(\gamma)) = \#_1(B(\gamma)) \qquad \textit{for } \gamma \in \mathcal{F}_{[m]}. \tag{3.5}$$

Assume that there is no set $S \subseteq [m]$ with $|S| = s$ where, for some i, $1 \le i \le t$, $A|_S$ contains F_i', $B|_S$ contains F_i'' or vice versa. Then $A =_c B$.

PROOF. We will use induction on the size of α and on m. There are two cases: $\alpha = \left[\begin{smallmatrix}0\\ \beta\end{smallmatrix}\right]$ or $\alpha = \left[\begin{smallmatrix}1\\ \beta\end{smallmatrix}\right]$. Assume $\alpha = \left[\begin{smallmatrix}0\\ \beta\end{smallmatrix}\right]$. Let C_1', C_1'' denote the matrices arising in (3.3) when we replace α by β, l by $l - 1$. We may decompose A,B to obtain

$$A =_c \begin{bmatrix} 00\cdots 0 & 11\cdots 1 \\ A_0 & A_1 \end{bmatrix}, \qquad B =_c \begin{bmatrix} 00\cdots 0 & 11\cdots 1 \\ B_0 & B_1 \end{bmatrix}. \tag{3.6}$$

Let γ be any $(0,1)$-column on rows $[m] - 1$, where γ has no submatrix β. Let

$$\gamma' = \begin{bmatrix} 0 \\ \gamma \end{bmatrix}, \qquad \gamma'' = \begin{bmatrix} 1 \\ \gamma \end{bmatrix}. \tag{3.7}$$

Now $\gamma', \gamma'' \in \mathcal{F}_{[m]}$ since they have no submatrix α and so

$$\#_1(A(\gamma')) = \#_1(B(\gamma')), \qquad \#_1(A(\gamma'') = \#_1(B(\gamma'')). \tag{3.8, 3.9}$$

But then we can deduce from (3.9) that $\#_1(A_1(\gamma)) = \#_1(B_1(\gamma))$ for all γ with no submatrix β and so from (3.8) $\#_1(A_0(\gamma)) = \#_1(B_0(\gamma))$ for all γ with no submatrix β. Applying induction, with α replaced by β, either there is some set S of rows of size $k + l - 1$ with $(A_0)|_S$ containing C_1', $(B_0)|_S$ containing C_1'' (or vice versa) or $A_0 =_c B_0$. In the former case note that in A and B, A_0 and B_0 are bordered with 0's and C_1' (resp. C_1'') bordered with 0's yields F_1' (resp. F_1'') so $A|_{S\cup\{1\}}$ contains F_i', $B|_{S\cup\{1\}}$ contains F_i'' (or vice versa). In the latter case we may delete from A and B those columns corresponding to A_0, B_0 and delete the first row to obtain matrices A_1, B_1 satisfying the hypothesis of this theorem and so by induction on m, $A_1 =_c B_1$, whence $A =_c B$.

Assume $\alpha = \left[\begin{smallmatrix}1\\ \beta\end{smallmatrix}\right]$ and begin, as in the other case, with (3.6), (3.7) and C_1', C_1'' arising in (3.3) when we replace α by β and k by $k - 1$. Now (3.8) holds for all γ' from a γ having no submatrix α, so $\#_1([A_0, A_1](\gamma)) = \#_1([B_0 B_1](\gamma))$ for all γ with no submatrix α. Using induction on m, we obtain $[A_0 A_1] =_c [B_0 B_1]$. Without loss of generality, we may assume that A and B have no columns in common and so $A_0 =_c B_1$, $A_1 =_c B_0$ (allowing matrices of no columns). Now (3.9) yields $\#_1(A_1(\gamma)) = \#_1(B_1(\gamma))$ for all γ with no submatrix β. Applying induction, with α replaced by β, either there is some set S of rows of size $k + l - 1$, where $(A_1)|_S$ contains C_1' and $(B_1)|_S$ contains C_1'' (or vice versa) or $A_1 =_c B_1$. In the former case we use $[A_0A_1] =_c [B_0B_1]$ to deduce that $(A_1)|_S$ and $(B_0)|_S$ have C_1' and $(A_0)|_S$ and $(B_1)|_S$ have C_1'' or vice versa. But by the bordering in (3.6) we obtain that $A|_{S\cup\{1\}}$ contains F_1' and $B|_{S\cup\{1\}}$ contains F_1'' (or vice versa). In the latter case we deduce that $A =_c B$ using $[A_0A_1] =_c [B_0B_1]$ etc. □

An alternate form of the theorem can be obtained by the same proof.

THEOREM 3.2. *Let A,B be two matrices on m rows satisfying* (3.6). *Assume that there is no set $S \in [m]$ with $|S| = s$ and, for some i with $1 \leq i \leq t$,*

$$A|_S \text{ contains } F_i' \text{ or } F_i''. \tag{3.10}$$

Then $A =_c B$. □

Either theorem can be thought of as encoding theorem, with the encoding given by the $A(\gamma)$'s. Next we have the linear algebra.

THEOREM 3.3. *Let A be a $(0,1)$-matrix on m rows with no submatrix which is a column permutation of $F \in \mathscr{F}$ (in* (3.5)). *Then the rank over $\mathbf{Q}$ of the space spanned by $A(\gamma)$ for $\gamma \in \mathscr{F}_{[m]}$ is equal to the number of distinct columns of A.*

PROOF. We may assume that A has n distinct columns since repeated columns can be deleted without affecting the rank. The rank is at most n, since the vectors are n-tuples. If rank $< n$, then the set of equations in n variables $x_1, x_2, \ldots, x_n$ given by

$$A(\gamma) \cdot (x_1, x_2, \ldots, x_n) = 0 \qquad \text{for all } \gamma \in \mathscr{F}_{[m]} \tag{3.11}$$

has a non-trivial integral solution $(e_1, e_2, \ldots, e_n)$. Let A_+ be the matrix with e_i copies of column i of A for $e_i > 0$ and let A_- have $-e_j$ copies of column j of A for $e_j < 0$. Then

$$\#_1(A_+(\gamma)) = \#_1(A_-(\gamma)) \qquad \text{for all } \gamma \in \mathscr{F}_{[m]} \tag{3.12}$$

and yet A_+ and A_- have no columns in common. Thus, by Theorem 3.1, there is an $S \subseteq [m]$, with $|S| = s$ and an i with $1 \leq i \leq t$ (of course, $i = 1$ here but not necessarily for examples in Section 4) with

$$A_+|_S \text{ contains } F_i', \quad A_-|_S \text{ contains } F_i'' \text{or vice versa.} \tag{3.13}$$

Since neither F_i' nor F_i'' have repeated columns, then $A|_S$ contains a column permutation of $[F_i'F_i''] \in \mathscr{F}$, a contradiction. Hence rank $= n$. □

In the Frankl and Pach proof of Theorem 1.1 [5], with $\mathscr{F}_{[m]}$ consisting of all columns of $k-1$ or fewer 1's, they refer to the pair A, $(e_1, e_2, \ldots, e_n)$ as a *null t-design*. We now obtain the bound.

THEOREM 3.4. *Let A be a $m \times n$ simple matrix with no submatrix which is a column permutation of $F \in \mathscr{F}$. Then*

$$n \leq |\mathscr{F}_{[m]}|. \tag{3.14}$$

PROOF. The rank in Theorem 3.3 is bounded by the number of vectors. □

THEOREM 3.5. *Let $A^{\mathscr{F}}$ be the $m \times |\mathscr{F}_{[m]}|$ simple matrix the columns of which are $\mathscr{F}_{[m]}$. Then A has no submatrix which is a column permutation of $F \in \mathscr{F}$.*

PROOF. This follows since α is a column of merge$(K_k, \vec{0}_l, \alpha)$ and $\mathscr{F}_{[m]}$ consists of those columns with no submatrix α. □

When the bound is exact in (3.14) we obtain a structural result, using $A^{\mathscr{F}}$.

THEOREM 3.6. *Let A be an $m \times |\mathscr{F}_{[m]}|$ simple matrix with no submatrix which is a column permutation of $F \in \mathscr{F}$. Then there exists a column permutation of A so that*

$$A \geqslant A^{\mathscr{F}}. \tag{3.15}$$

PROOF. Use P. Hall's theorem. If k columns of $A^{\mathscr{F}}$ are not covered by k columns of A then, since the columns of $A^{\mathscr{F}}$ are $\mathscr{F}_{[m]}$, the rank of the space spanned by $A(\gamma)$ for $\gamma \in \mathscr{F}_{[m]}$ will be less than $|\mathscr{F}_{[m]}|$, violating Theorem 3.3. □

It is true for $\mathscr{F} = \text{merge}(K_k, \vec{0}_l, \alpha)$ that the bound in Theorem 3.4 follows from Theorem 1.1. However, the other results provide additional information on the consequences of forbidding merge$(K_k, \vec{0}_l, \alpha)$. It is interesting to find that $A^{\mathscr{F}}$ is the unique minimal matrix among extremal matrices avoiding merge$(K_k, \vec{0}_l, \alpha)$. Is there a shifting proof as in Alon's result [1]? The sequence of theorems that we have followed will be used in Section 4 with different $\mathscr{F}_{[m]}$.

4. MORE EXAMPLES

The linear algebra results of Section 3 adapt to other $\mathscr{F}_{[m]}$. Our first example demonstrates the method. Let

$$\mathscr{F}_{[m]} = \{\gamma \ m \times 1(0, 1)\text{-column} \mid \gamma \text{ has at most two 1's or exactly 3 conservative 1's}\} \tag{4.1}$$

whence

$$|\mathscr{F}_{[m]}| = \binom{m}{2} + \binom{m}{1} + \binom{m}{0} + m - 2. \tag{4.2}$$

The proof of Theorem 3.1 for this $\mathscr{F}_{[m]}$ will yield $\mathscr{F}$ from:

$$F_1' = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad F_1'' = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix},$$

$$F_2' = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad F_2'' = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix},$$

$$F_3' = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad F_3'' = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}, \tag{4.3}$$

$$F_4' = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \qquad F_4'' = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix},$$

$$F_5' = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \qquad F_5'' = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix},$$

$$F_6' = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}, \qquad F_6'' = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

THEOREM 4.1. *Theorems* 3.1, 3.2, 3.3, 3.4, 3.5 *and* 3.6 *hold with* $\mathcal{F}_{[m]}$ *as in* (4.1). $\mathcal{F}$ *as in* (3.4), F_i', F_i'' *as in* (4.3) *and* $t = 6$, $s = 4$.

PROOF. Theorem 3.1 requires most of our work. The result is easy for $m < 4$. For $m = 4$ we note that there are four special columns,

$$\gamma_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \gamma_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad \gamma_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \gamma_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{4.4}$$

where

$$\#_1(A(\gamma)) = \#_1(B(\gamma)) \qquad \text{for all } 4 \times 1 \ \gamma \text{ with } \gamma \neq \gamma_1, \gamma_2, \gamma_3. \tag{4.5}$$

Assume that A, B have no columns in common. A case argument will yield that A has an F_i' and B has an F_i'', or vice versa. For example, assume $\#_1(A(\gamma_1)) = c > 0$. Then A has c copies of γ_1. Hence $\#_1(B(\gamma_1)) = 0$. Using (4.5) with the two γ with 3 consecutive 1's, we deduce that B has c copies of both columns of 3 consecutive 1's. Now using $\gamma = (0, 1, 1, 0)^{\mathrm{T}}$ in (4.5) we deduce that A has c copies of $(0,1,1,0)^{\mathrm{T}}$. Since $\#_1(A(\gamma_4)) \geq c$, B has either γ_2, γ_3 or γ_4. If, for example B has $d > 0$ copies of γ_2 then A has d copies of $(1,0,1,0)^{\mathrm{T}}$ and $(0,0,1,1)^{\mathrm{T}}$. Thus $\#_1(A((0,0,1,0)^{\mathrm{T}})) = c + d$ and so B has d copies of $(0,0,1,0)^{\mathrm{T}}$, yielding F_1' in A and F_1'' in B. The remaining cases arise similarly. Note that $A = F_i'$ and $B = F_i''$ yield examples satisfying (4.5).

For $m > 4$, we use induction on m. Let $\mathcal{F}_{[m]-\{i\}}$ denote the columns as given in (4.1) with m replaced by $m - 1$, where we index the rows $1, 2, \ldots, m$ with i missing. We may delete row 1 and apply induction, since any column in $\mathcal{F}_{[m]-\{1\}}$ with a 0 added as row 1 is in $\mathcal{F}_{[m]}$. So $A|_{[m]-\{1\}} =_c B|_{[m]-\{1\}}$. Similarly, we may delete row m and apply induction to obtain $A|_{[m]-\{m\}} =_c B|_{[m]-\{m\}}$.

We may delete row 2 and apply induction if we can verify $\#_1(A(\gamma)) = \#_1(B(\gamma))$ for the $m \times 1$ γ having exactly three 1's in rows 1, 3 and 4. However, this follows from deleting row m. So $A|_{[m]-\{2\}} =_c B|_{[m]-\{2\}}$. We may delete row 3 and apply induction if we can verify $\#_1(A(\gamma)) = \#_1(B(\gamma))$ for two $m \times 1$ γ, the first having 1's only in rows 1, 2 and 4 and the second having 1's only in rows 2, 4 and 5. The first case follows by deleting row m and the second follows by deleting row 1. Thus $A|_{[m]-\{3\}} =_c B|_{[m]-\{3\}}$.

Now use Lemma 4.2 below to obtain that, for $T=\{1,2,3,m\}$,

$$A|_T \text{ contains } E_4,\ B|_T \text{ contains } O_4 \text{ or vice versa.} \tag{4.6}$$

Thus $A|_T$ contains F_1' and $B|_T$ contains F_1'' or vice versa. This contradiction finishes the proof of Theorem 3.1. Note that the columns $\mathcal{F}_{[m]}$ are precisely those columns with no submatrix γ_1, γ_2, γ_3 in (4.4) and yet each $[F_i'F_i'']$ has such a column. Thus Theorem 3.5 follows. □

The arguments will generalize to $\mathcal{F}_{[m]}$ consisting of all columns of at most k 1's and columns with $k+1$ 1's appearing consecutively. The following Lemma encapsulates a proof idea of Ryser [7].

Lemma 4.2. *Let A,B be two matrices on m rows and let $S\subset[m]$ and $|S|=s$ have the property that*

$$A|_{[m]-\{i\}}=_c B|_{[m]-\{i\}} \qquad \textit{for all } i\in S. \tag{4.7}$$

Assume that A and B have no common columns. Then

$$A|_S \textit{ contains } E_s, \qquad B|_S \textit{ contains } O_s, \qquad \textit{or vice versa.} \tag{4.8}$$

Proof. Without loss of generality, we may assume that $S=\{1,2,\ldots,s\}$, since row permutations of E_s and O_s are just column permutations. Decompose A by grouping columns based on their initial s entries as follows

$$A=_c\begin{bmatrix} \vec{0}_s\vec{0}_s\cdots\vec{0}_s & \cdots & \alpha\alpha\cdots\alpha & \cdots & \vec{1}_s\vec{1}_s\cdots\vec{1}_s \\ A_{\vec{0}_s} & \cdots & A_\alpha & \cdots & A_{\vec{1}_s}\end{bmatrix}, \tag{4.9}$$

where α is any $s\times 1$ (0, 1)-column and A_α may have no columns. Decompose B similarly. Let e_i be the $s\times 1$ (0, 1)-column with a single 1 in row i. Let $\oplus$ denote modulo 2 sum by entries (addition in $GF(2)^s$). By (4.7), deleting row i leaves matrices equal apart from a column permutation, and so examining their first $s-1$ entries we obtain

$$[A_\alpha A_{\alpha\oplus e_i}]=_c[B_\alpha B_{\alpha\oplus e_i}] \tag{4.10}$$

for each $s\times 1$ (0, 1)-column α. Moreover since A,B have no common columns,

$$A_\alpha=_c B_{\alpha\oplus e_i}. \tag{4.11}$$

This is true for each i and $\{e_1,e_2,\ldots,e_s\}$ form a basis for $GF(2)^s$, so we obtain

$$A_\alpha=_c A_\beta=_c B_\gamma=_c B_\delta \tag{4.12}$$

for all $s\times 1$, α, β, γ, δ with α,β having an even number of 1's, γ, δ having an odd number of 1's, or vice versa. If $A_{\vec{0}_s}$ has columns, we obtain that

$$A|_S \text{ contains } E_s, \qquad B|_S \text{ contains } O_s, \tag{4.13}$$

and if $A_{\vec{e}_1}$ has columns, we obtain that

$$A|_S \text{ contains } O_s, \qquad B|_S \text{ contains } E_s. \tag{4.14}$$

If neither holds then we deduce that A, and hence B, has no columns, a contradiction. □

Our second example relates to Corollary 2.3, with β being a given $(k+l)\times 1$ (0, 1)-columns of k's. We define

$$\mathscr{F}=\{\text{merge}(K_k, \alpha, \beta) \mid \alpha \text{ any } l\times 1 \ (0,1)\text{-column}\} \tag{4.15}$$

and $\mathscr{F}_{[m]}$ will be the columns with no submatrices merge$(\vec{1}_k, \alpha, \beta)$ for all $l\times 1$ (0, 1)-vectors α. To construct $\mathscr{F}_{[m]}$ for $m\geq k+l-1$, note that k 1's of β split the l 0's of β into $k+1$ blocks (some possibly empty) and let β_i be the number of 0's in the ith block. We have $\beta_1+\beta_2+\cdots+\beta_{k+1}=l$. Now from an $(m-l)\times(1$ (0, 1)-vector δ of $k-1$ 1's, we form an $m\times 1$ vector δ^* of $k+l-1$ 1's by inserting β_1 1's in the first rows, β_2 1's after the first 1 of δ, β_3 1's after the second 1 of δ, etc., β_k 1's after the $(k-1)$st 1 of δ, and β_{k+1} 1's in the final rows. You may envisage δ^* as consisting of $k+1$ *blocks* of 1's of sizes β_1, β_2+1, $\beta_3+1,\ldots,\beta_k+1$, β_{k+1}. We define

$$\mathscr{F}_{[m]}=\{\gamma\ m\times 1\ (0, 1\text{-column} \mid \gamma\leq\delta^* \text{ for some } (m-l)\times 1 \text{ column } \delta \text{ of } k-1 \text{ 1's}\}. \tag{4.16}$$

Note that the columns δ^* are the *maximal* columns in $\mathscr{F}_{[m]}$. We deduce that

$$|\mathscr{F}_{[m]}|=2^l\left(\binom{m-l}{k-1}+\binom{m-l}{k-2}+\cdots+\binom{m-l}{0}\right). \tag{4.17}$$

To see this, note that for any column $\gamma\in\mathscr{F}_{[m]}$ we can uniquely identify up to $k-1$ *leading* 1's and exactly *l additional rows* possibly with 1's and all other rows with 0's as follows. Imagine the 0th leading 1 to be in row 0. The ith leading 1 must leave $\beta_{i+1}+\beta_{i+2}+\cdots+\beta_{k+1}$ rows after it. The l additional rows are placed, much as in δ^*, in the β_i rows after the $(i-1)$st leading 1 and, where p is the final leading 1, the final $\beta_{p+2}+\beta_{p+3}+\cdots+\beta_{k+1}$ rows are additional rows.

Let $\alpha_1, \alpha_2, \ldots, \alpha_{2^l}$ be the 2^l $l\times 1$ (0, 1)-vectors. Then define, for $1\leq i\leq 2^l$,

$$F_i'=\text{merge}(E_k, \alpha_i, \beta), \qquad F_i''=\text{merge}(O_k, \alpha_i, \beta). \tag{4.18}$$

THEOREM 4.3. *Theorems* 3.1, 3.2, 3.3, 3.4, 3.5 *and* 3.6 *hold with* $\mathscr{F}_{[m]}$ *as in* (4.16), $\mathscr{F}$ *as in* (4.15) *with* F_i', F_i'' *given in* (4.18), $t=2^l$, *and* $s=k+l$.

PROOF. Theorem 3.1 requires most of our work. The result is easy for $m<k+l$. Assume that $m=k+l$. Assume that A,B have no common columns. By construction, the maximal columns δ^* of $\mathscr{F}_{[m]}$ will have a single 0 in a row where β has a 1. Thus

$$\#_1(A(\gamma))=\#_1(B(\gamma)) \tag{4.19}$$

for all γ with at least one 0 where β has a 1. A column γ not having this property can be written $\gamma=\text{merge}(\vec{1}_k, \alpha, \beta)$ for some $l\times 1$ α. Among all such γ with $\#_1(A(\gamma))\neq\#_1(B(\gamma))$, choose one with α maximal, say merge$(\vec{1}_k, \bar{\alpha}, \beta)$. But then (4.19) holds for

$$\gamma=\text{merge}(\mu, \alpha, \beta),\ \mu \text{ any } k\times 1(0,1)\text{-column},\ \alpha \text{ any } l\times 1\ (0,1)\text{-column}$$
$$\text{with } \bar{\alpha}\leq\alpha,\ \bar{\alpha}\neq\alpha. \tag{4.20}$$

Since A,B have no common columns we deduce that A, B have no column γ as in (4.20), since if $\hat{\gamma}$ was the maximal such column in either A or B, say $\hat{\gamma}$ is in A, then

$\#_1(A(\hat{\gamma})) = \#_1(B(\hat{\gamma}))$ would imply, by maximality of $\hat{\gamma}$ and $\bar{\alpha}$, that $\hat{\gamma}$ is in B, a contradiction. We also have that (4.19) holds for

$$\gamma = \text{merge}(\mu, \bar{\alpha}, \beta),\ \mu \text{ any } k \times 1 \text{ (0, 1)-column},\ \mu \neq \bar{1}_k. \tag{4.21}$$

Now let A', B' be obtained from A, B by selecting those columns ρ so that $\rho|_S = \bar{\alpha}$, where S is the set of rows where β has 0's. Then, for γ as in (4.21), $\#_1(A(\gamma)) = \#_1(A'(\gamma))$ and $\#_1(B(\gamma)) = \#_1(B'(\gamma))$. Let A'', B'' be obtained from A', B' by deleting rows indexed by S. Then (recalling $m = k + l$) we have $\#_1(A''(\gamma)) = \#_1(B''(\gamma))$ for all $k \times 1$ γ with $\gamma \neq \bar{1}_k$. We note that A'', B'' have no columns in common, so by Lemma 4.2 (or Proposition 2.5, [2]) we obtain that

$$A'' \text{ contains } E_k,\ B'' \text{ contains } O_k \text{ or vice versa} \tag{4.22}$$

and so, by construction of A'', B''.

$$A \text{ contains merge}(E_k, \bar{\alpha}, \beta),\ B \text{ contains merge}(O_k, \bar{\alpha}, \beta) \text{ or vice versa.} \tag{4.23}$$

For $m > k + l$, we use induction on m. Let $\mathcal{F}_{[m]-\{i\}}$ denote the columns as given in (4.16) with m replaced by $m - 1$, where the rows are indexed $1, 2, \ldots, m$ with row i missing. We need to show that

$$\#_1(A|_{[m]-\{i\}}(\gamma)) = \#_1(B|_{[m]-\{i\}}(\gamma)) \qquad \text{for } \gamma \in \mathcal{F}_{[m]-\{i\}} \tag{4.24}$$

for $k + l$ different rows i, so that we may conclude by induction that $A|_{[m]-\{i\}} =_c B|_{[m]-\{i\}}$. We will do this first for rows

$$1 + \beta_1,\ 1 + \beta_1 + (\beta_2 + 1),\ 1 + \beta_1 + (\beta_2 + 1) + (\beta_3 + 1), \ldots, 1 + \beta_1 + (\beta_2 + 1) + \cdots + (\beta_k + 1). \tag{4.25}$$

Recall that the maximal columns δ^* can be constructed from $k + 1$ blocks of 1's, the sizes of which, in order from the top, are β_1, $\beta_2 + 1$, $\beta_3 + 1, \ldots, \beta_k + 1$, β_{k+1} where the first and last blocks occupy the first and last rows respectively. Thus we can show (4.24) for $i = 1 + \beta_1$ since any maximal column in $\mathcal{F}_{[m]-\{1+\beta_1\}}$, when a zero is inserted in row $1 + \beta_1$, yields a maximal column in $\mathcal{F}_{[m]}$. Thus, by induction,

$$A|_{[m]-\{1+\beta_1\}} =_c B|_{[m]-\{1+\beta_1\}}. \tag{4.26}$$

In turn, we can show (4.24) for $i = 1 + \beta_1 + (\beta_2 + 1)$. For a maximal column γ in $\mathcal{F}_{[m]-\{1+\beta_1+(\beta_2+1)\}}$, let γ' be the column obtained by inserting a 0 in row $1 + \beta_1 + (\beta_2 + 1)$. Either $\gamma' \in \mathcal{F}_{[m]}$ or $1 + \beta_1 + (\beta_2 + 1)$ must split the second blocks in γ, and so γ' must have a 0 in row $1 + \beta_1$. But then, by (4.26),

$$\#_1(A(\mu)) = \#_1(B(\mu)) \qquad \text{for all } \mu \leq \gamma'. \tag{4.27}$$

Now we can obtain (4.24) for $i = 1 + \beta_1 + (\beta_2 + 1)$ by induction. We continue in this way. We can show (4.24) for $i = 1 + \beta_1 + (\beta_2 + 1) + \cdots + (\beta_j + 1)$ $(j \leq k)$ since if γ is a maximal column in $\mathcal{F}_{[m]-\{i\}}$ and γ' is the column obtained by inserting a 0 in row $1 + \beta_1 + (\beta_2 + 1) + \cdots + (\beta_j + 1)$, then either γ' is a maximal column in $\mathcal{F}_{[m]}$ or the 0 added must split some block b for $2 \leq b \leq j$. In the latter case the first i blocks are not all as high as they can be, and so there is a 0 in one of the rows $1 + \beta_1$, $1 + \beta_1 + (\beta_2 + 1), \ldots, 1 + \beta_1 + (\beta_2 + 1) + \cdots + (\beta_{j-1} + 1)$. Apply the previously obtained results for the appropriate row to obtain (4.24).

By flipping the analysis, we see that we obtain (4.24) for $i = m - \beta_{k+1}$, $m - \beta_{k+1} - (\beta_k + 1), \ldots, m - \beta_{k+1} - (\beta_k + 1) - \cdots - (\beta_2 + 1)$. We now take an arbitrary row r and establish (4.24) for $i = r$. Let γ be a maximal column in $\mathcal{F}_{[m]-\{r\}}$ and let γ' be γ with a 0 inserted in row r. Either $\gamma' \in \mathcal{F}_{[m]}$ or the 0 splits a block of 1's of γ, say the

pth block. Either the blocks $1, 2, \ldots, p$ are not all as high as they could be, in which case γ' has a 0 in one of rows $1+\beta_1$, $1+\beta_1+(1+\beta_2)$, etc. and then we obtain (4.24); or the blocks $p, p+1, \ldots, k+1$ are not all as low as they could be, in which case γ' has a 0 in one of rows $m-\beta_{k+1}$, $m-\beta_{k+1}-(\beta_k+1)$, etc. and then we obtain (4.24). So (4.24) holds for $i=r$. We can now use Lemma 4.2 to deduce that $A =_c B$.

To show that $A^{\mathcal{F}}$ satisfies Theorem 3.5 we need to show that each $\gamma \in \mathcal{F}_{[m]}$ has no submatrix merge$(\vec{1}_k, \alpha, \beta)$ for any $l \times 1$ α. Then, since merge$(\vec{1}_k, \alpha, \beta)$ is a column of merge(K_k, α, β), we will be done. To see that merge$(\vec{1}_k, \alpha, \beta)$ is not a submatrix, review the leading 1's analysis given after (4.19) to note that they can be identified with the 1's of $\vec{1}_k$ in merge$(\vec{1}_k, \alpha, \beta)$, but that there are only $k-1$ leading 1s. $\square$

One can obtain other results using different column-based measures. Also, one can generalize $A(\gamma)$ appropriately to obtain a linear algebra proof of Alon's result [1], a question raised in [3]. Let γ be an appropriate integral vector and A an integral $m \times n$ matrix with $\gamma, A \geqslant 0$. Then $A(\gamma)$ is a $1 \times n$ (0, 1)-row with a 1 in column j if, for each $i \in [m]$, column j of A either has its ith entry equal to 1 *or* to the ith entry of γ. If A is a (0, 1)-matrix, this definition agrees with what we have used above.

Acknowledgement

This research was supported by NSERC.

References

1. N. Alon, On the density of sets of vectors, *Discr. Math.*, **46** (1983), 199–202.
2. R. P. Anstee, General forbidden configuration theorems, *J. Combin. Theory, Ser. A*, **40** (1985), 108–124.
3. R. P. Anstee, A forbidden configuration theorem of Alon, *J. Combin. Theory, Ser. A*, **47** (1986), 16–27.
4. R. P. Anstee and Z. Füredi, Forbidden submatrices, *Discr. Math.*, **62** (1986) 225–243.
5. P. Frankl and J. Pach, On the number of sets in a null t-design, *Europ. J. Combin.*, **4** (1983), 21–23.
6. Z. Füredi, private communication.
7. H. J. Ryser, Combinatorial configurations, *SIAM J. Appl. Math.*, **17** (1969), 596–602.
8. N. Sauer, On the density of families of sets, *J. Combin. Theory, Ser. A*, **13** (1972), 145–147.
9. S. Shelah, A combinatorial problem: stability and order for models and theories in infinitary languages, *Pac. J. Math.*, **4** (1972), 247–261.
10. V. N. Vapnik and A. Ya. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Theory Prob. Applics*, **16** (1971), 264–280.

Received 20 January 1991 and accepted 14 March 1995

R. P. Anstee
Department of Mathematics,
University of British Columbia,
#121-1984 Mathematics Road,
Vancouver, B.C., Canada V6T 1Z2
and
Department of Mathematics and Statistics
University of Otago, P.O. Box 56, Dunedin, New Zealand